


~~MEMORANDUM FOR:~~ Deputy Assistant Director for Central Reference

THROUGH: Assistant Director for Operations

SUBJECT: Machine Translation Evaluation Test

1. In response to your request of 13 July 1959 for an opinion sampling of machine translation products, I am enclosing six copies of a report which describes the design, results, and some observations of the MT product evaluation test administered on a limited scale in Foreign Documents Division on 15 July 1959.

2. This test, as constructed and administered, represents an introductory attempt at developing a method for evaluating machine translation products, and I hope it might contribute toward the development of a more appropriate and sophisticated statistical method of evaluation test procedure.


J. J. BAGHAZ
Chief, Foreign Documents Division, CG

Enclosure:
Machine Translation Evaluation
Test (six copies)

Distribution:
Orig & 1 - Addressee
✓ 1 - AD/O
1 - BS/FDD

15 July 1959

FOREIGN DOCUMENTS DIVISION
MACHINE TRANSLATION EVALUATION TEST

OBJECTIVE

The objective of the test was to determine the relative quality of two machine translation products as considered in the subjective judgment of raters experienced in Russian language scientific and technical translation.

TEST DESIGN

Four sample translations of the same Russian text were arrayed in columns identified only as A, B, C, D on one sheet of paper (See Tab "A"). Column A represented the 8 June 1959 machine printout of GAT. Column B was introduced as a control translation to bridge the obvious quality gap between the two machine translations and the human criterion translation, thus eliminating the possibility of overwhelming bias for the criterion translation. Column C represents the 20 November 1958 machine printout of CMF. Column D represents the human criterion translation. During instruction to the judges no mention was made of the fact that any of the four sample translations were other than machine translations.

Rather than submit the sample translations to the 16 judges with a blanket request to rank them in order of general excellence of translation as was done in a similar experiment on 9 May 1959, it was decided to formulate a series of statements which would apply to certain qualitative factors desirable in a translation. Thus, eight statements reflecting three significant levels of translation quality--meaning transfer (statements 1, 2), scanning suitability (statements 3,4), publication suitability (statements 5, 6, 7, 8)--were devised. The judges were asked to rate the translations in rank order from most to least for each statement as applicable. In Part II the judges were asked to correct vocabulary and grammar and were provided with the original Russian text for reference. (See Tab "A")

RESULTS

The following general tabulation gives the rank order (most to least) of each translation by statement. Table 1 in Tab "B" presents raw score information for each translation by rank order for each statement.

1. Readability	B D C A
2. Meaning comprehension	B D C A
3. Scanning without pre-editing	B D C A
4. Scanning with pre-editing	A C B
5. Rigorous editing, frequent referral to original text	A C D B
6. Editing, minimum referral to original text	B D C A
7. Normal editing, no referral	B D C A
8. Suitable for publication	D B C A

In general, the combination EDCA--control, criterion, CMT, GAT-- was preferred by the raters with respect to the statements (1, 2, 3, 6, 7, 8) reflecting positive qualities desired in a translation. Table 2 (Tab "B") shows the frequency of rank order combinations for each statement.

Since the primary object of the test was to determine which of the two machine translations (A or C) is considered of better quality, the rank order of the human translations (B and D) was excluded from any further consideration.

Table 3 (Tab "B") presents the raw scores and proportions of CA and AC choices with respect to the total rating population. The proportion is actually a mean and can be regarded as an average probability. Further computation of measures of variability and reliability was not performed. No attempt was made to establish the correlation between the various statements. However, cursory examination of grouped data points to a high degree of correlation between statements 1 and 2. Thus, if a text is readable, it is usually comprehensible.

OBSERVATIONS

The following general observations are based on an examination of the grouped data in Tables 1, 2 and 3:

1. CMT was considered both more readable and comprehensible than GAT.

2. CMT could be more easily scanned without pre-editing.

3. GAT requires more frequent referral to the original foreign language text than CMT.

4. An insignificant number felt that CMT could be submitted to normal editing.

5. An even less significant number than in (4) felt that CMT was suitable for publication more so than GAT.

Thus, the results of the test indicate that CMT is considered to be of generally better translation quality than GAT.

The vocabulary and grammatical corrections performed in Part II were not subjected to analysis because of a lack of sufficient number of samples (half of the sample population) and the apparent rater disinterest in actively prosecuting this part of the test.

The reactions were studied between the ethyl ester of pyrocatecholophosphorous acid and triaryl bromomethanes. Upon the interaction of the indicated compounds there are formed the pyrocatechol esters of triarylmethylphosphinic acids. They are obtained upon hydrolysis of the latter by weak hydrochloric acid pyrocatechol and triarylmethylphosphinic acids.

By us there were studied in the present investigation the reactions between esters the mixed of phosphorous acid, a type /formula/ where R plus C₂H₅/5/, and triaryl bromomethanes. The reaction between the ethylpyrocatechol ester of phosphorous acid and triaryl bromomethanes by analogy is must go with alkylphosphorous esters at a reaction.. /formula/

Experimental given showed, that a reaction actually proceeds according to the indicated equation. Thus, for example, upon the heating of a mixture of the triaryl bromomethane and the ethylpyrocatechol ester of phosphorous acid there occur the evolution of ethyl bromide and the formation of a crystalline substance, which is the pyrocatechol ester of a triarylmethylphosphinic acid. The reaction of hydrolysis for the establishment of the structure of a obtained compound was conducted by diluted hydrochloric acid at heating from 180 to 200° in sealed tubes. With the product of hydrolysis it are pyrocatechol and a triarylmethylphosphinic acid.

Reactions between the ethyl ester of pyrocatechol phosphorous acid and triaryl bromomethanes were studied. The pyrocatechol esters of triarylmethylphosphinic acids are formed by the interaction of the compounds indicated. Pyrocatechol and triarylmethylphosphinic acids are obtained by hydrolysis of the latter with weak hydrochloric acid.

The reaction between mixed esters of phosphorous acid of the type C₆H₄-O₂-POR (where R = C₂H₅) and triaryl bromomethanes were studied by us in the present investigation. The reaction between the ethylpyrocatechol ester of phosphorous acid and triaryl bromomethanes, by analogy with alkylphosphorous esters, should go according to the reactions: (cf. p 1479)

The experimental data indicated that the reaction, in fact, proceeds according to the above-mentioned equation. Thus, for example, upon heating of a mixture of triaryl bromomethane and the ethylpyrocatechol ester of phosphorous acid, there occurs the separation out of ethyl bromide and the formation of a crystalline substance which is the pyrocatechol ester of triarylmethylphosphinic acid. For determining the structure of the obtained compounds, hydrolysis with weak hydrochloric acid at 180-200 deg in sealed tubes was conducted. The products of hydrolysis were pyrocatechol and triarylmethylphosphinic acid.

Reactions between ethyl ester of pyrocatechol phosphorous acid and with triaryl bromomethanes were studied. Upon interaction of above-mentioned compounds the pyrocatechol esters of triarylmethylphosphinic acids are formed. Upon the saponification of the latter with weak hydrochloric acid pyrocatechol and the triarylmethylphosphinic acids were obtained.

In the present investigation reactions between the mixed esters of phosphorous acid were studied, type refer p 1479 pg 1 --- (where R eq C₂H₅), and with triaryl bromomethanes. reaction between ethylpyrocatechol ester of phosphorous acid and with triaryl bromomethanes by analogy with alkylphosphorous esters must proceed according to the reaction = refer p 1479 pg 2 .

The experimental data did show, that reaction actually proceeds according to the above-mentioned equation. Thus, for example, upon the heating of the mixture of the triaryl bromomethane and ethylpyrocatechol ester of phosphorous acid isolation of bromide ethyl and formation of crystalline substance occurs, which is pyrocatechol ester of triarylmethylphosphinic acid. For the establishment of the structure of obtained compound was the reaction of the saponification with diluted hydrochloric acid upon the heating from 180 to 200 deg in the sealed tubes carried out. The products of the saponification are pyrocatechol and the triarylmethylphosphinic acid.

Reactions between the ethyl ester of pyrocatechol phosphorous acid and triaryl bromomethanes were studied. Upon interaction of the indicated compounds, the pyrocatechol esters of triarylmethylphosphinic acids are formed. Upon hydrolysis of the latter with weak hydrochloric acid, pyrocatechol and triarylmethylphosphinic acids are obtained.

In the present investigation, we studied the reaction between mixed esters of phosphorous acid of the type C₆H₄-O₂-POR (where R = C₂H₅) and triaryl bromomethanes. The reaction between the ethylpyrocatechol ester of phosphorous acid and triaryl bromomethanes, by analogy with alkylphosphorous esters, should go according to the reactions: (schematics omitted: tr.)

The experimental data showed that the reaction indeed proceeds according to the indicated equation. Thus, for example, upon heating of a mixture of triaryl bromomethane and the ethylpyrocatechol ester of phosphorous acid, there occurs the separation out of ethyl bromide and the formation of a crystalline substance which is the pyrocatechol ester of triarylmethylphosphinic acid. For determination of the structure of the produced compounds, they were hydrolyzed with weak hydrochloric acid at 180-200° in sealed tubes. The products of the hydrolysis were pyrocatechol and triarylmethylphosphinic acid.

769

MACHINE TRANSLATION EVALUATION TEST

Part I

Rank each translation with respect to the followings:

	1	2	3	4
1. Readability				
2. Meaning comprehension				
3. Suitable for scanning as to subject matter <u>without editing</u>				
4. Suitable for scanning as to subject matter <u>only after editing</u>				
5. Would require rigorous editing with frequent referral to original foreign language text				
6. Would require editing with minimum referral to original foreign language text				
7. Would require normal editing (no referral to original foreign language text)				
8. Suitable for publication (no editing required)				

Part II

Referring to the original Russian text, examine each translation.
CIRCLE vocabulary errors and UNDERLINE grammatical errors.

Изучены реакции между этиловым эфиром пирокатехин-фосфористой кислоты и триарилбромметанами. При взаимодействии указанных соединений образуются пирокатехиновые эфиры триарилметилфосфиновых кислот. При омылении последних слабой соляной кислотой получены пирокатехин и триарилметилфосфиновые кислоты.

В настоящем исследовании нами изучались реакции между смешанными эфирами фосфористой кислоты, типа $A_2 \dots \dots \dots$ и триарилбромметанами. Реакция между этилпирокатехиновым эфиром фосфористой кислоты и триарилбромметанами по аналогии с алкилфосфористыми эфирами должна идти по реакции: $A_2 \dots \dots \dots$

Экспериментальные данные показали, что реакция действительно протекает по указанному уравнению. Так, например, при нагревании смеси триарилбромметана и этилпирокатехинового эфира фосфористой кислоты происходит выделение бромистого этила и образование кристаллического вещества, представляющего собой пирокатехиновый эфир триарилметилфосфиновой кислоты. Для установления строения полученного соединения была проведена реакция омыления разбавленной соляной кислотой при нагревании от 180 до 200° в запаянных трубках. Продуктом омыления являются пирокатехин и триарилметилфосфиновая кислота.

Tab "B"

Table 1

Raw Scores by Rank Order for Each Statement

1. Readability

	1	2	3	4
A			2	(13)
B	(11)	4		
C			(13)	2
D	5	(11)		

Result: B D C A

2. Meaning
Comprehension

	1	2	3	4
A			1	(12)
B	(12)	3		
C		1	(12)	
D	5	(11)		

Result: B D C A

3. Scanning without pre-editing

	1	2	3	4
A			3	(5)
B	(12)	3	1	
C		3	7	1
D	6	(10)		

Result: B D C A

4. Scanning with pre-editing

	1	2	3	4
A	⑧	5		
B			③	2
C	6	⑦		
D			2	2

Result: A C B BD

5. Rigorous editing, frequent referral to original

	1	2	3	4
A	⑩	4		
B			1	③
C	5	⑧		
D			③	1

Result: A C D B

6. Editing, minimum referral to original

	1	2	3	4
A	3	1	2	③
B	⑦	2		
C	4	2	⑤	1
D		⑤		

Result: B D C A

7. Normal editing, no referral to original

	1	2	3	4
A			1	③
B	⑧	5		
C			④	1
D	6	⑦		

Result: B D C A

8. Suitable for publication

	1	2	3	4
A				③
B	4	③		
C			③	
D	⑦	2		

Result: D B C A

Tab "B"

Table 2

Frequency of Rank Order Combinations For Each Statement

1. Readability

BDCA	9
DECA	4
BDAC	2
D	1

2. Meaning Comprehension

BDCA	9
DECA	4
BD	2
D	1

3. Scanning without pre-editing

BDCA	5
BD	4
DECA	2
EDC	2
DBAC	1
DCB	1
DB	1

4. Scanning with pre-editing

CA	4
AC	3
ACDB	2
ACED	1
CABD	1
ACB	1
A	1
C	1
No score	2

5. Rigorous editing, frequent referral to original

AC	5
CA	3
ACDB	2
A	2
ACED	1
CADB	1
C	1
No score	1

6. Editing, minimum referral

C	3
BDCA	2
BDAC	1
DEC	1
DECA	1
BCA	1
CA	1
BD	1
BC	1
D	1
A	1
No score	2

7. Normal editing, no referral

DB	4
BD	3
BDCA	2
DECA	1
BDAC	1
BDC	1
D	1
B	1
No score	2

8. Suitable for publication

D	4
B	2
BDCA	2
DB	2
DECA	1

Tab "B"

Table 3
Scores and Proportions of Comparative Judgments

	STATEMENT	CA		AC	
		SCORE	P	SCORE	P
MEANING (Readability)	1	13	.81	2	.13
	2	13	.81	0	.00
SCANNING	3	7	.44	1	.06
	4	6	.38	8	.50
PUBLICATION	5	5	.31	10	.63
	6	9	.56	2	.13
	7	4	.25	1	.06
	8	3	.19	0	.00

N = 16